

# **THE NCEP GLOBAL ENSEMBLE FORECASTING SYSTEM**

Zoltan Toth<sup>1</sup>

National Centers for Environmental Prediction  
1 GSC (Beltsville, MD) at NCEP  
W/NP20, World Weather Building,  
5200 Auth Road  
Camp Springs, MD 20746, USA

## **SUMMARY**

Global ensemble forecasts have been produced as part of the NCEP operational suite since December 1992. Different ensemble based products have been generated and distributed, through various channels, to a wide range of users both nationally and internationally. Evaluation of the quality of the products indicate that the ensemble forecasts can provide substantial economic value, beyond that provided by the higher resolution control forecast, for a wide range of users.

## **INTRODUCTION**

Ensemble forecasting is the only practical technique to assess the flow dependent predictability of the atmosphere, and to create probabilistic forecasts reflecting it (Ehrendorfer, 1997). Though statistical techniques can also produce probabilistic forecasts based on traditional single control Numerical Weather Prediction (NWP) model forecasts, it has been demonstrated that such guidance has a markedly inferior performance (Talagrand, 1999, personal communication; Toth et al., 1998).

Ensemble forecasting entails running an NWP model (or several model variants) a number of times, with slightly perturbed initial conditions, to assess the forecast uncertainty due to errors in the initial conditions and possibly in model formulation. Ensemble forecasting became a routine practice at major operational NWP centers around the world. After it was implemented at NCEP

in December 1992, ECMWF (Molteni et al, 1996), FNMOC (Rennick, 1995), the Canadian Meteorological Center (CMC, Houtekamer et al., 1996), the Japan Meteorological Agency (JMA, Kobayashi et al., 1996), and the South African Weather Bureau (SAWB, Tennant, 1996, personal communication) also implemented ensemble forecasting while other centers are considering its implementation.

## **RATIONALE**

Ensemble forecasting is based on the recognition that the atmosphere is a chaotic system in which any error in initial condition or model formulation leads to loss of predictability after a finite period of time. It is also found that depending on the flow, predictability greatly varies in time and space (high vs. low predictability example). There are correspondingly large variations in the "hit rate", or "success rate" of forecasts. Ensemble forecasting is to identify these flow dependent variations in advance, at the time the forecasts are made. These variations in the expected success rate of the forecasts, if known in advance, can greatly enhance the information content of the weather forecasts (ensemble based forecasts of variations in predictability).

## **GOAL**

The ultimate goal of ensemble forecasting is to provide flow dependent full probability forecasts (in the form of full and joint probability distributions) for all atmospheric variables. It has been shown that when forecasts are expressed in the form of full probability distributions their information content is greatly enhanced as compared to a traditional, two-level (dichotomous, corresponding to yes-no outcomes) probability forecasts (Toth et al., 1998, information content). Note that the goal of ensemble forecasting is markedly different and broader than that of a single control forecast, which is to provide only a best estimate (or the expected state) of the atmosphere.

## **TECHNIQUE**

Ensemble forecasting entails multiple model integrations, started with slightly perturbed initial conditions, possibly using different model versions to also account for model related uncertainty.

**a) Initial perturbations.** There are three main techniques used at the different centers.

(1) *Breeding* (NCEP, FNMOC, SAWB, JMA). This is a technique that identifies those possible analysis errors that can amplify most rapidly (Toth and Kalnay, 1993; 1997)

(2) *Singular Vectors* (SVs, ECMWF, JMA). This is a technique to identify perturbation structures that can grow fastest in the forecasts. SVs should, but in practice have not been, constrained by the probability at which different initial perturbation patterns can occur in the analysis. Note that JMA tested both the breeding and the SV methods and decided to use the breeding method in the future (Takano, 1999, personal communication).

(3) *Perturbed observations technique* (CMC). For each perturbed analysis, a

separate analysis cycle is run, where all observations are perturbed with random noise representing the error in the observations.

Methods (1) and (3) are similar that they attempt to capture patterns that can occur in the analysis as errors. Method (1), however, ignores perturbation patterns that initially are not growing. Methods (1) and (2) are similar in that they both ignore non-growing patterns (i. e., they both use only dynamically conditioned perturbations); method (1), however, disregards transient perturbation growth and captures only perturbations that can sustain their growth at the perturbation amplitudes representative of analysis errors. All three techniques have merit and their comparative evaluation is a subject of ongoing research. Method (1) is by far the simplest and computationally least demanding of the three methods.

The breeding method was modified for ensemble applications so that the regional rescaling of the initial perturbations reflect the estimated uncertainty present in analyses (Toth and Kalnay, 1997).

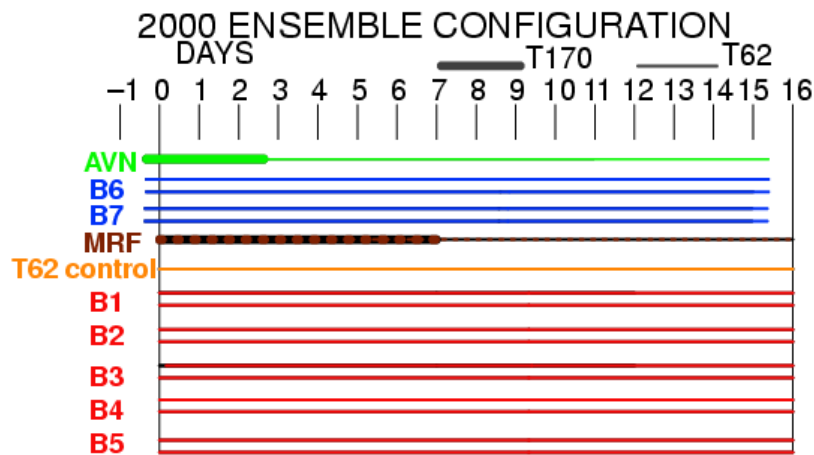
**b) Model perturbations.** When ensemble forecasting was first implemented at major NWP operational centers (Molteni et al., 1996; Toth and Kalnay, 1993) it was designed to assess forecast uncertainty related to errors present in the initial conditions. The initial errors project on atmospheric instabilities and amplify in time, rendering forecasts, even if we had a perfect model, useless beyond a finite period of time (Lorenz, 1969). In practice, however forecast uncertainty also arises due to the fact that we use simplified numerical models to predict the behavior of the atmosphere. The use of such models lead to the emergence of errors, in addition to those due to inaccurate initial conditions.

Part of the overall error due to model imperfectness can be classified as systematic, and another part as random or stochastic. We can define the systematic part of the model related error as that which can be reproduced if the model is run many times over similar cases. In practice, these errors can only be estimated using finite verification statistics. Systematic errors are due to inaccurate model formulation, such as inadequate parametrization of certain subgridscale processes.

The stochastic part of the error is not reproducible because it does not depend on the flow regime (except possibly in a statistical sense). Stochastic errors arise at each integration time step due to numerical inaccuracies, the use of finite truncation, and other inaccuracies that act in a random fashion. The stochastic errors, just like the initial errors, turn in time into the direction of fastest growing perturbation directions, increasing errors associated with atmospheric instabilities.

There are different attempts at accounting for model related uncertainty in ensemble forecasting. At CMC (Houtekamer et al, 1996; Houtekamer and Lefaivre, 1997), several versions of an NWP model are developed and used in parallel with each other. These versions possibly differ from each other in horizontal resolution, treatment of orography, convection and radiation parametrization, etc. For each ensemble model integration started with unique and slightly different initial conditions, a different model version is used. The goal is to capture systematic differences or errors in model forecasts, though the real atmospheric solution still differs more from the ensemble members than the individual forecasts from each other.

At ECMWF (Buizza et al., 1999), after each time step within a model integration, stochastic multiplicative noise is added to the diabatic forcing term. After the forcing from all parametrized



processes is added up, the net forcing is multiplied by a number chosen randomly in the [0.5, 1.5] interval, making the impact of the complete physics package stochastic. The goal is to represent the inherent uncertainty in the parametrization of subgrid-scale processes that leads to the emergence of stochastic errors during model integrations.

The NCEP ensemble forecasting system does not account for model related errors yet. Postprocessing probabilistic forecasts based on the ensemble, however, is effective in creating directly usable output by eliminating biases in probabilities.

## CURRENT CONFIGURATION

The current operational ensemble configuration consists of running:

At 0000 UTC each day:

- T170 high resolution control out to 7 days, after which this run gets truncated and is run out to 16 days at a T62 resolution
- T62 control that is started with a truncated T170 analysis
- 10 perturbed forecasts each run at T62 horizontal resolution. The perturbations are from five independent breeding cycles (see Toth and Kalnay, 1993, 1997). Each perturbation is both added to, and subtracted from the control analysis.

At 1200 UTC:

- T170 control out to 3 days ("Aviation forecast"), that gets truncated and run at T62 resolution out to 16 days
- Two pairs of perturbed forecasts based on two independent breeding cycles (four perturbed integrations out to 16 days.)

## PRODUCTS

A number of products are being generated based on the NCEP global ensemble forecasts. The list of products has been expanding and still covers only a fraction of possibly useful guidance that can be made using the ensemble forecasts (see also

[http://sgi62.wwb.noaa.gov:8080/ens/info/ens\\_detbak.html#ENSEMBLE PRODUCTS](http://sgi62.wwb.noaa.gov:8080/ens/info/ens_detbak.html#ENSEMBLE_PRODUCTS)).

**Ensemble mean.** This is the most basic forecast guidance from the ensemble. Due to the

ensemble's ability to filter out unpredictable events, this field gives a better estimate for the expected value of the future state of the atmosphere. Note that because the unpredictable, often smaller scale events are selectively filtered out, this field is smoother than any of the individual forecasts. It is therefore essential to consider other information from the ensemble, like ensemble spread and/or single contour plots, along with the ensemble mean, that can reveal the variability exhibited by the ensemble members that contribute to the mean.

**Ensemble spread.** The standard deviation around the ensemble mean is considered another basic guidance product, indicating the variance of ensemble members around the mean.

**Normalized ensemble spread.** The ensemble spread here is expressed in terms of a ratio of the actual ensemble spread over the ensemble spread averaged for the given lead time over the preceeding month. It is for the detection of anomalously high or low spread (indicating low or high predictability, respectively), irrespective of lead time and geographical location. Current and recent forecast plots for the ensemble mean, spread, and normalized spread are available on the web.

**Single contour (spaghetti) diagrams.** A selected contour level of a given variable is plotted on the same figure for each individual ensemble member. It provides a quick overview of all ensemble forecasts. Note that in areas of small gradients, large differences in the spaghetti lines may occur, without the ensemble members being substantially different. Examples of single contour plots are available on the web (see September 1999 cases).

**Cluster means or tubes.** These are statistically derived products that attempt to capture prevailing and important aspects of the ensemble forecasts (Tracton and, 1993; Atger, 1999). Their primary purpose is to condense information and they should not be considered more than alternative ways of representing the forecasts. The notes made for ensemble mean forecasts are also relevant for cluster means.

**Probabilistic forecasts.** This is considered the most important and comprehensive product based on an ensemble. For any given weather event that need to be predicted, the number of ensemble forecasts indicating that event is counted. The ratio of forecasts predicting the event, over the total number of forecasts is the relative forecast frequency that can be interpreted as a probabilistic forecast. Current and recent Probabilistic Quantitative Precipitation Forecasts (PQPF) are available on the web.

## POST PROCESSING

**Biases in probabilities.** The quality of the ensemble forecasts is compromised by errors both in model formulation (systematic model errors or biases) and ensemble techniques (lack of adequate representation of model errors). In particular, for the lack of adequate representation of model (and not initial value) related uncertainty in ensembles, the spread of the NCEP (and other) ensemble forecasts is insufficient at longer lead times. This leads to probabilistic forecasts that, over the long run, do not match corresponding observed frequency values. This problem can be easily addressed by a simple calibration process (Zhu et al., 1996). The calibrated probabilistic forecasts are very reliable, i. e., events that are predicted with say, a 60% probability occur, over the long run, at 60% of the time. It is important to emphasize that this performance is achieved despite the fact that model uncertainties are not yet accounted for in the NCEP ensemble.

**Model systematic errors.** Before probabilistic forecasts are made for sensible weather elements,

the individual ensemble forecasts can be statistically postprocessed to reduce possible systematic errors or model biases. Statistical postprocessing has also been a critical element in the interpretation of traditional single control forecasts (e. g., Carter et al., 1989). Note that the purpose of statistical postprocessing of the ensemble forecasts is different from that of a single control forecast. MOS, for example, not only attempts to eliminate the bias from the forecasts on which it is applied but also hedges the forecasts toward climatology (the larger the expected forecast error, the more so). A single control forecast is normally used to provide a best estimate of the future state of the atmosphere, and hedging serves well this purpose. Ensemble forecasting, however, has a different goal, providing a full forecast probability distribution. In this case hedging, that brings all forecasts, intended to represent the inherent forecast uncertainty, closer to climatology is counterintuitive.

## **VERIFICATION**

Objective evaluation of an ensemble involves the generation of a host of statistics, including ensemble mean rms errors, ensemble spread (which should ideally match the ensemble mean error), analysis rank histograms (Talagrand diagrams), Brier Skill Score (BSS), Ranked Probability Skill Score (RPSS), Relative Operating Characteristics (ROC), and Information Content (IC) (Zhu et al., 1996). The most important measures are those evaluating the performance of probabilistic forecasts (BSS, RPSS, ROC, IC). Basically, probabilistic forecasts have to meet two criteria to be of value: (1) they need to be reliable (or consistent with observations), i. e., events predicted with a given probability should verify with a frequency of the given forecast probability; and (2) they need to have resolution, i. e., have to be as different from climatological frequencies as possible (preferably close to 0 and 1 probability values). The best probabilistic system would give a probability of 1 for events that actually occur, and 0 for all other possible events. Because the atmosphere is chaotic, it is usually not possible to achieve this theoretical limit of skill. The skill scores listed above reward probabilistic forecast systems that approach this theoretical limit by being both reliable and exhibiting high resolution. The quality of ensemble forecasts based on the NCEP system was compared to those based on the ECMWF operational system. It was found that the NCEP ensemble forecasts exhibit higher scores for the first couple of days of integration, while the ECMWF ensemble forecasts have higher scores beyond that (Talagrand, 1999, personal communication; Zhu et al., 1996). This is probably due to the use of more realistic initial perturbations in the NCEP system, while a slightly higher quality forecast model in the ECMWF system.

The performance of the ensemble forecast system can also be compared to that of a higher resolution control forecast. These two systems use approximately the same amount of computational resources. Toth et al. (1998) found that the ensemble system was superior in all measures beyond 72 hours lead time.

## **ECONOMIC VALUE**

The ultimate test of the quality of a forecast system is made through an analysis of the economic benefit different users can gain from using it. The economic benefit associated with the use of an ensemble of forecasts vs. a higher resolution control forecast can also be compared. A simple decision making model can be used where all potential users of weather forecasts are characterized by the ratio between the cost of their action to prevent weather related damages, and the loss that they incur in case they do not protect their operations. As Mylne (1999),

Richardson (2000), and Toth et al. (2000).

Toth et al. (2000) showed, in cases of appreciable forecast uncertainty (after 24-72 hours lead time on the synoptic scales) the ensemble forecast system can be used by a much wider range of users, and with significantly greater economic benefits, than the higher resolution control forecast. This confirms results with more traditional verification measures. The added benefit of the ensemble approach derive from (1) the ensemble's ability to differentiate between high and low predictability cases, and (2) the fact that it provides a full forecast probability distribution, allowing the users to tailor their weather forecast related actions to their particular cost/loss situation.

## USAGE

The ensemble forecasts serve multipurpose applications, various ranges-variables-properties, by providing forecast probability distributions of the atmosphere:

### a) Variance-Covariance information

6-12 hr: To be used in analysis (planned)

### b) General forecast guidance

24-72 hr: Short-range applications, currently underutilized

Boundary conditions for a Limited Area Ensemble (Tracton & Du)

72-168 hr: Medium-range guidance - most used

8-14 day: Extended-range guidance (CPC)

### c) Tropical depressions/storms

72-168 hr: Early warning of possible developments

### d) Time/space evolution of error variance

24-168 hr: Targeted observations

The NCEP global ensemble forecasts are used extensively by HPC and CPC within NCEP, and are widely used by NWS field offices. Beyond the NWS, the users of ensemble forecasts are thematically and geographically widely distributed, including:

USER GROUP	BASE	AREA OF INTEREST	PRODUCTS	INTEREST
US Air Force	US	Global	PQPF and others	Aviation, etc.
US Forest Service	US	Western US	PQPF and others	Fire weather
Hydrological agencies	US, Central and South America	US, Central and South America, Africa	PQPF, temperature	Flood mitigation
Energy companies	US, Europe	US, Europe	Height, temperature, PQPF	Fuel delivery planning
Weather derivative industry	US	US	Height, etc	Predictability of weather

## **DISTRIBUTION**

The NCEP global ensemble forecasts are distributed through the following channels:

a) NCEP/EMC Web page

It offers an array of graphical products, including ensemble mean, spread, normalized spread, Probabilistic Quantitative Precipitation (PQPF), and single contour charts (currently not available). It is also a central source for information on different aspects of the ensemble system, including other distribution channels.

b) NWS/OSO, and NCEP/ftp ftp servers. These servers contain conveniently arranged "enspost" files, for easy downloading of ensemble data for 20 or so individual variables, and postprocessed information (e. g., PQPF data).

c) AWIPS Satellite broadcast system. 500 hPa height, 850 hPa temperature, mean sea level pressure, and accumulated precipitation data will be distributed effective when the ensemble gains full operational status. Graphical products are also planned to be distributed. Note that special processing and display software needs to be developed for the AWIPS platforms for their optimum use of the ensemble data.

d) Graphics in NAWIPS metafile format are available to the NCEP centers.

e) Outside distribution links. NOAA/OAR/CDC in Boulder offers graphical products on its web page and serves as an archive for past ensemble forecast data

f) As part of a research agreement with ECMWF, NCEP and ECMWF exchanges their ensemble forecast data on a daily basis. Note that similar exchanges are planned with CMC and FNMOC.

## **RECENT CHANGES**

Recent changes to the NCEP ensemble forecast system include:

Effective 12UTC April 6 1999

Increase in initial perturbation amplitude size

Effective 07 December 1998 at 12 UTC :

Change in regional rescaling procedure for setting initial perturbation amplitudes

Effective 06 May 1998 at 12UTC:

New seasonally varying analysis uncertainty estimates introduced into regional rescaling procedure

Effective March 1997:

Ensemble forecast data are available on OSO server

Effective February 11 1997:

Ensemble precipitation forecast data made available

## **PERSONNEL**

The following people have contributed substantially in the past to the global ensemble developmental work:

Eugenia Kalnay - Technique development (University of Maryland)

Tim Marchok - Graphics (SAIC)

Currently the following people work on global ensemble related projects:



NAME	AFFILIATION	AREA	% TIME DEVOTED	COMMENTS
Istvan Szunyogh	UCAR Visiting Scientist	Technique development	67	
Yuejian Zhu	GSC at EMC	Verification	50	
Richard Wobus	GSC at EMC	Postprocessing	100	
Zoltan Toth	GSC at EMC	Coordination, technique development	67	

## SHORTCOMINGS/PLANS

SYMPTOMS	PROBLEM	SOLUTION	IMPLEMENTATION DATE	TPB
Suboptimal performance in terms of systematic and random errors	Low Horizontal Resolution	Increase Horizontal Resolution	Dependent on computer upgrades	
		T126 for first 84 hours	April 2000	TPB/2000/04
		T126 for first 7.5 days	January 2001	
<b>Inability To:</b>	Too Few Ensemble Members	Increase Ensemble Membership	Ongoing	
Identify extreme/rare events				
Provide users with with very high or low cost-loss ratios		Introduce 6 more perturbed forecasts at the 12 UTC cycle	April 2000	TPB/2000/04
Provide adequate guidance for targeted observations and analysis applications for reliable covariance estimates				
Provide boundary conditions for Limited Area Ensembles 2-4 x per day		Introduce 10 perturbed forecasts at both 06 & 18 UTC cycles	January 2001	
Too large spatial variations in initial perturbations	Breeding cycle is 24 hours long	Change breeding cycle length to 6 hours	January 2001	
Initial perturbation size does not reflect changes in data coverage	Use of climatologically fixed perturbation amplitudes in breeding	Make rescaling procedures in breeding adaptive by incorporating information on data coverage/observation errors from analysis	2001	
Insufficient perturbation amplitude at medium-extended ranges; cloud of ensemble does not encompass verification*	Stochastic and systematic model errors are not accounted for	Combine ensembles, after bias correction from many centers and create a multi-model ensemble	2001	
		Develop a model that properly accounts for stochastic and systematic errors	2002-2003	
Lack of sufficient forecast guidance products*	Ensemble forecasts are not postprocessed extensively	Introduce bias correction for the first and second moments of the ensemble	2000	TPB/2000/04
		Express forecasts in terms of anomalies with reference to reanalysis climatology	2001	
		Provide probability forecasts for stations based on ensemble anomaly guidance	2002	

\*denotes collaborative effort is needed

## **CREDITS**

The development and operational implementation of the ensemble forecasting system would not have been possible without the efforts of a number of people, including:

Steve Tracton, Mark Iredell, Suranjana Saha, Hua-Lu Pan, Stephen Lord (EMC), Masao Kanamitsu (CPC), Joe Irwin, Maxine Brown, Cliff Dye, and Joe Johnson (NCO)

## **REFERENCES**

Atger, F., 1999. Tubing: an alternative to clustering for the classification of ensemble forecasts. *Weather and Forecasting*, 14, 5, 741-757.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic simulation of model uncertainty in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 125, 2887-2908.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Wea. Forecasting*, 4, 401-412.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, 124, 1225-1242.

Houtekamer, P. L., and L. Lefaivre, 1997: Using ensemble forecasts for model validation. *Mon. Wea. Rev.*, 125, 2416-2426.

C. Kobayashi, C., K. Yoshimatsu, S. Maeda, and K. Takano, 1996: Dynamical one-month forecasting at JMA. Preprints of the 11th AMS Conference on Numerical Weather Prediction, Aug. 19-23, 1996, Norfolk, Virginia, 13-14.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, 21, 289-307.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaigis, 1996: The ECMWF ensemble system: Methodology and validation. *Q. J. R. Meteorol. Soc.*, 122, 73-119.

Mylne, K.R., 1999 The use of forecast value calculations for optimal decision making using probability forecasts. Preprints of the 17th AMS Conference on Weather Analysis and Forecasting, 13-17 September 1999, Denver, Colorado, 235-239.

Rennick, M. A., 1995: The ensemble forecast system (EFS). Models Department Technical Note 2-95, Fleet Numerical Meteorology and Oceanography Center. p. 19. [Available from: Models Department, FLENUMMETOCCEN, 7 Grace Hopper Ave., Monterey, CA 93943.]

Richardson, D. S., 2000a: Skill and economic value of the ECMWF ensemble prediction system, *Q.J.R.Meteorol. Soc.*, 126, 649-668.

Toth, Z., and E. Kalnay, 1993: Ensemble Forecasting at the N MC: The generation of perturbations. *Bull. Amer. Meteorol. Soc.*, 74, 2317-2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, 125, 3297-3319.

Toth, Z., Y. Zhu, T. Marchok, . Tracton, and E. Kalnay, 1998: Verification of the NCEP global ensemble forecasts. *Preprints of the 12th Conference on Numerical Weather Prediction*, 11-16 January 1998, Phoenix, Arizona, 286-289.

Tracton, M. S. and E. Kalnay, 1993: Ensemble forecasting at NMC: Operational implementation. *Wea. Forecasting*, 8, 379-398.

Zhu, Y., G. Iyengar, Z. Toth, S. M. Tracton and T. Marchok, 1996: Objective evaluation of the NCEP global ensemble forecasting system . *Preprints* , 15th AMS Conference on Weather Analysis and Forecasting, Norfolk, Virginia.